IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Weissman et al.	Art Unit:	2162
Serial No.:	10/741,303	Examiner:	Dennis Y. Myint
Filed:	December 18, 2003	Conf. No.:	4367
Title:	METHODS AND SYSTEMS FOR DETECTING AND EXTRACTING INFORMATION		

Mail Stop Appeal Brief - Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

BRIEF ON APPEAL

This Brief on Appeal perfects the Notice of Appeal filed April 9, 2008 and appeals the rejections set forth in the Office action mailed January 10, 2008.

(1) Real Party in Interest

This case is assigned of record to Google Inc., which is the real party in interest.

(2) Related Appeals and Interferences

There are no known related appeals and/or interferences.

(3) Status of Claims

Claims 1-29 and 35-48 are pending.

Claims 1-29 and 35-48 are under consideration.

Claims 30-34 have been canceled.

Claims 1-29 and 35-48 stand rejected.

Claim 1, 15, and 47 are in independent form.

Claims 1, 15, and 47 are involved directly in the appeal. Claims 2-14, 16-29, 35-46, and 48 are not directly involved in the appeal but rather are involved only by virtue of their dependency from one or more of claims 1, 15, and 47.

(4) Status of Amendments

A response pursuant to 37 C.F.R. § 1.116 was filed on March 4, 2008. The Advisory Action mailed March 24, 2008 indicated that this response would be entered for purposes of appeal. Accordingly, all amendments have been entered.

(5) Summary of Claimed Subject Matter

Search engines are mechanisms that allow users to search vast numbers of documents for information. *See, e.g., specification*, page 1, line 23-24. Commonly, search engines provide a query field on a user interface into which a user can enter one or more "search terms." *See, e.g., specification*, page 1, line 24-page 2, line 1. Typically, search engines match these search terms to terms in the text of the documents. *See, e.g., specification*, page 2, line 8-9. More complex methods of searching also exist. *See, e.g., specification*, page 2, line 18. However, most of these methods are also term-based. *See, e.g., specification*, page 2, line 18-19. For example, advanced Boolean searches or advanced fuzzy text searches are both generally based on terms. *See, e.g., specification*, page 2, line 19-20.

With term-based searches, the document's terms are compared with the defined search terms or with similar terms having similar rankings. *See, e.g., specification*, page 2, line 20-22. Problems with such term-based searches arise due to the multiple meanings of words, the use of synonyms, and difficulties in determining the importance of terms if the terms are not used frequently. *See, e.g., specification*, page 3, line 1-3. Further, it is difficult to search for a concept using term-based searching methods. *See, e.g., specification*, page 3, line 3-4. For example, if it is desired to search all documents that mention the concept "California cities", each city must be enumerated as a search term. *See, e.g., specification*, page 3, line 4-6.

To address these and other issues, the present inventors have developed systems and techniques in which requests for information from a user include definitions of concept lists.

Claim 1 relates to a computer-implemented method. The method includes:

receiving, from a user, a request for information that includes a definition of a concept list (*See, e.g., specification*, page 7, line 16-21; page 12, line 4-6; page 15, line 18-19) comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept (*See, e.g., specification*, page 8, line 1-2; page 11, line 7-10; page 12, line 6-9; page 12, line 11-22), and a target scope that characterizes a document region to which the concept list is to be applied (*See, e.g., specification*, page 8, line 5-11; page 10, line 10-12; page 13, line 3-11; page 13, line 12-14);

receiving a definition of an extraction rule, wherein the extraction rule definition comprises an extraction scope that characterizes a document region to be extracted (*See, e.g., specification*, page 17, line 19-page 18, line 1);

determining a target score for the document regions of the article, wherein the score represents how well the document regions relate to the concept list (*See, e.g., specification*, page 8, line 21-page 9, line 2; page 11, line 4-6; page 13, line 14-17; page 16, line 4-page 17, line 11);

applying the extraction rule to the article to determine an extract from the article, wherein the application of the extraction rule is based on the determined target score (*See, e.g., specification*, page 13, line 14-17; page 18, line 5-11; page 19, line 10-20); and

outputting the extract in response to the request for information (*See, e.g., specification*, page 9, line 14; page 20, line 10-13).

Claim 15 relates to an article comprising one or more computer-readable data storage media containing program code operable to cause one or more machines to perform operations *See, e.g., specification*, page 6, line 14-page 7, line 2; page 5, line 2-20. The operations include:

receiving, from a user, a request for information that includes a definition of a concept list (*See, e.g., specification*, page 7, line 16-21; page 12, line 4-6; page 15, line 18-19) comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept (*See, e.g., specification*, page 8, line 1-2; page 11, line 7-10; page 12, line 6-9; page 12, line 11-22), and a target scope that characterizes a document region to which the target rule is to be applied (*See, e.g., specification*, page 8, line 5-11; page 10, line 10-12; page 13, line 3-11; page 13, line 12-14);

receiving a definition of an extraction rule, wherein the extraction rule definition comprises an extraction scope that characterizes a document region to be extracted (*See, e.g., specification*, page 17, line 19-page 18, line 1);

determining a target score for the document regions of the article, wherein the score represents how well the document regions relate to the concept list (*See, e.g., specification*, page 8, line 21-page 9, line 2; page 11, line 4-6; page 13, line 14-17; page 16, line 4-page 17, line 11);

applying the extraction rule to the article to determine an extract from the article, wherein the application of the extraction rule is based on the determined target score (*See, e.g., specification*, page 13, line 14-17; page 18, line 5-11; page 19, line 10-20); and

outputting the extract in response to the request for information (*See, e.g., specification*, page 9, line 14; page 20, line 10-13).

Claim 47 relates to a computer-implemented method for extracting a subset of a document. The method includes:

receiving, from a user, a request for information that describes a combination of two or more concept lists (*See, e.g., specification*, page 7, line 16-21; page 12, line 4-6; page 15, line 18-19), wherein each concept list is defined by an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept (*See, e.g., specification*, page 8, line 1-2; page 11, line 7-10; page 12, line 6-9; page 12, line 11-22), wherein the two or more concept lists are combined using an operation to define a target definition that is to be detected (*See, e.g., specification*, page 7, line 21-page 8, line 1; page 11, line 6-7; page 12, line 10-11; page 13, line 10-12; page 13, line 20-22);

receiving a description of a document region targeted for extraction (*See, e.g., specification*, page 8, line 5-11; page 10, line 10-12; page 13, line 3-11; page 13, line 12-14; page 17, line 19-page 18, line 1);

accessing a document (*See, e.g., specification*, page 8, line 12; page 18, line 12-17);

based on the target definition and the document regions targeted for extraction, extracting one or more regions of the accessed document (*See, e.g., specification*, page 8, line 21-page 9, line 14; page 10, line 16-18; page 18, line 5-6; page 18, line 18-page 20, line 9); and

making the extracted regions available for output in response to the request for information (*See, e.g., specification*, page 9, line 14; page 20, line 10-13).

(6) Grounds of Rejection to be Reviewed on Appeal

As set forth in the following concise statements, the following grounds for rejection are presented for review on appeal:

Ground 1: Whether claims 1-4, 8, 12-13, 15-18, 22, 26-27, 35-39, 42-43, and 44 are properly rejected under 35 U.S.C. § 103(a) as obvious over U.S. Patent No. 5,724,571 to Woods (hereinafter "Woods") and U.S. Patent Publication No. 2003/0115191 to Copperman et al. (hereinafter "Copperman").

Ground 2: Whether claims 9, 23, 40, 41, 45, 46, 47, and 48 are properly rejected under 35 U.S.C. § 103(a) as obvious over Woods, Copperman, and U.S. Patent Publication No. 6,763,349 to Sacco (hereinafter "Sacco").

(7) Argument

The organization of the arguments presented hereinafter follows the organization of the grounds for rejection to be reviewed on appeal set forth above. In particular, a separate boldfaced heading for each ground presented for review follows.

Objections to the Specification

The specification is objected to as failing to provide proper antecedent basis for "computer-readable data storage media" recited in independent claim 15 and its dependencies. The objection cites to 37 C.F.R. § 1.75(d)(1) and M.P.E.P. § 608.01(o).

For the sake of convenience, 37 C.F.R. § 1.75(d)(1) is now reproduced.

> "(d)(1)The claim or claims must conform to the invention as set forth in the remainder of the specification and the terms and phrases used in the claims must find clear support or antecedent basis in the description so that the meaning of the terms in the claims may be ascertainable by reference to the description." *See* 37 C.F.R. § 1.75(d)(1) (emphasis added).

As shown, the specification is not objectionable based solely on a failure to provide *in ipsis verbis* antecedent basis for the terms and phrases used in the claims. Instead, the specification can provide clear support for such terms and phrases and satisfy the requirements of 37 C.F.R. § 1.75(d)(1).

Against this backdrop, attention is respectfully directed to para. [0012] of the specification, which is now reproduced for the sake of convenience.

> "Embodiments of computer-readable media include, but are not limited to, an electronic, optical, magnetic, or other storage or transmission device capable of providing a processor, such as the processor in communication with a touch-sensitive input device, with computer-readable instructions. Other examples of suitable media include, but are not limited to, a floppy disk, CD-ROM, magnetic disk, memory chip, ROM, RAM, an ASIC, a configured processor, all optical media, all magnetic tape or other magnetic media, or any other medium from which a computer processor can read instructions." *See Specification*, para. [0012] (emphasis added).

The specification thus explicitly describes that computer-readable media include storage devices capable of providing a processor with computer-readable instructions, and moreover lists several examples thereof. Both clear support and antecedent basis for the "computer-readable data storage media" recited in claim 15 and its dependencies are provided in the specification and the requirements of 37 C.F.R. § 1.75(d)(1) satisfied.

When this was pointed out in the response file March 4, 2008, the Advisory Action mailed March 24, 2008 contended that:

> "'computer readable media' and 'computer-readable data storage media' are not the same. Applicant's attention is respectfully directed to the fact that 'Computer-readable data storage media' is not explicitly defined..." *See Advisory Action mailed March 24, 2008*, page 2, second paragraph.

Applicant respectfully disagrees. To begin with, as discussed above, the specification explicitly describes that computer-readable media include storage devices that provide computer-readable instructions and gives examples thereof.

Moreover, to the extent that the Advisory Action can be read as contending that the specification is somehow deficient for failing to explicitly define "computer-readable data storage media," applicant respectfully disagrees. Applicants are not under any burden to explicitly define all the terms that appear in a claim. Rather, claims terms are presumed to have their ordinary and customary meaning unless they are defined otherwise. *See, e.g., M.P.E.P.* § 2111.01 Plain Meaning I. THE WORDS OF A CLAIM MUST BE GIVEN THEIR "PLAIN MEANING" UNLESS THEY ARE DEFINED IN THE SPECIFICATION (emphasis added).

Accordingly, the specification is not deficient for any failure to explicitly define "computer-readable data storage media."

Applicant respectfully requests that the objection to the specification be withdrawn.

Ground 1: Rejections under 35 U.S.C. § 103(a) over Woods and Copperman

Claims 1 and 15 were rejected under 35 U.S.C. § 103(a) as obvious over Woods and Copperman.

As discussed above, claim 1 relates to a computer-implemented method that includes receiving, from a user, a request for information that includes a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept, and a target scope that characterizes a document region to which the concept list is to be applied, receiving a definition of an extraction rule, determining a target score for the document regions of the article, applying the extraction rule to the article to determine an extract from the article, and outputting the extract in response to the request for information. The extraction rule definition comprises an extraction scope that characterizes a document region to be extracted. The score represents how well the document regions relate to the concept list. The application of the extraction rule is based on the determined target score.

Claim 15 relates to an article that includes one or more computer-readable data storage media containing program code operable to cause one or more machines to perform operations. The operations are related to the activities recited in claim 1.

The rejections of claims 1 and 15 gloss over a fundamental distinction between the recited subject matter and the content of Woods and Copperman. In particular, claims 1 and 15 relate to requests for information and responses to such requests. The requests for information include a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept, and a target scope that characterizes a document region to which the concept list is to be applied. Thus, the claims encompass the situation where a user submits a search query defined by concepts, and the relationship between those concepts, rather than by a mere collection of terms.

In contrast, requests for information in Woods and Copperman do not include these features. For example, Woods' requests for information are understood to be a single "search query phrase (consisting of one to many terms)." *See Woods*, col. 5, line 67-col. 6, line 3. *See also Woods*, col. 3, line 28-35 (describing his system as "particularly effective" at handling short (i.e., two to six word) search queries).

In Copperman, the requests for information can be developed during an iterative, guided search process. In particular, Copperman describes an iterative process in which a search query is received and concepts/topics are matched to the search query. *See, e.g., Copperman*, para. [0054]. *See also id.*, para. [0051] (describing that terms in a user query that are evidence of concepts are first extracted and then used as the basis for guided search routines). Copperman's FIGS. 9A-9E illustrate one example of the matching of concepts/topics to a search query in Copperman's iterative, guided search process. *See, e.g., id.*, paras. [0081], [0017]-[0021]. As

shown in, e.g., FIG. 9A, a textual search query is first received from a user. *See, e.g., id.*, para. [0081]. As shown in, e.g., FIG. 9C, the response to the search query can include results that match a "primary group feature" that was "spotted" in the terms of the search query, as well as "related features." As best understood, Copperman's features are terms or phrases. *See id.*, para. [0052]. The "related features" can be used to narrow the scope of the search and reduce the number of documents "in play." *See, e.g., id.*, paras. [0054], [0081]. Thus, the requests for information in Copperman are an initial search query along with any subsequently selected "related features" that narrow the scope of the initial search query.

Therefore, neither Woods' single search query phrase, nor Copperman's initial search query and any subsequently selected "related features," describe or suggest requests for information that include a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept, and a target scope that characterizes a document region to which the concept list is to be applied, as recited in claims 1 and 15. Lacking such requests, Woods and Copperman also fail to describe or suggest responding to such requests, as recited in claims 1 and 15.

When this was pointed out in the response file March 4, 2008, the Advisory Action mailed March 24, 2008 contended that:

> "'receiving, from a user, a request for information that includes a definition of a concept list' (Woods, Figure 4, i.e.. Input search query 410; Woods, Column 5 line 67 through Column 6 line 1, i.e., a search query phrase (consisting of one to many terms) is input)..." *See Advisory Action mailed March 24, 2008*, page 2, fourth paragraph.

Thus, as best understood, the rejection is based on the contention that a multi-term search query is a "definition of a concept list," as recited in claims 1 and 15. *See also Advisory Action mailed March 24, 2008*, page 3, line 1 (contending that "[a] search query phrase that consists one

to many terms is a list of concept"); *Office action mailed January 10, 2008*, page 4, line 1-4 (contending that Woods' receipt of a "*search query phrase* ***(consisting of one to many terms)***" constitutes the receipt of a request for information that includes a definition of a concept list) (emphasis in original).

Applicant respectfully disagrees and submits that a multi-term search query is not a definition of a concept list. For example, the multi-term search queries "Washington monument," "Detroit Red Wings," "Honda Accord," "Yul Brynner," and "U.S. Patent and Trademark Office" do not define a concept list. The instant application carefully distinguishes term-based searching, like that in Woods and that brings with it many problems, from the concept-based searching recited in the pending claims. *See, e.g., specification*, page 1, line 1 – page 3, line 6. *See also* § (5) Summary of Claimed Subject Matter, *supra*. As a result, the Office cannot conflate term-based searching with concept-based searching.

Moreover, Woods himself is directed to the proposition that traditional term-based searching is insufficient as failing to encompass concepts. *See, e.g., Woods*, col. 1, line 15-42. For example, Woods describes that term-based searching can generate an excessively large number of hits, omit paraphrase variations from a result set, and ineffectively rank passages in a result set. *Id.*

Indeed, Woods describes that a semantic network of terms and concepts (i.e., term/concept relationship network 110) is to be constructed *independently of a query* by analyzing a corpus of documents. *See, e.g., Woods*, col. 5, line 7-14. The semantic network is then used to connect terms in queries with the text of the documents in the corpus. *See, e.g., Woods*, col. 5, line 7-14. *See also id.*, col. 5, line 66-col. 7, line 56 (describing how the terms in a search query are matched to the classified documents). Thus, Woods himself does not consider

a collection of terms in a search query to be a list of concepts. Instead, the terms in a search query must be matched with the concepts in a semantic network so that concept-based searching can proceed.

Since the contention that a multi-term search query forms a "definition of a concept list" is inconsistent both with applicant's specification and the cited Woods reference, it is not reasonable. Accordingly, an obviousness rejection on this basis cannot be sustained.

Copperman does not remedy these deficiencies in Woods. As discussed above, the requests for information in Copperman are an initial search query along with any subsequently selected "related features" that narrow the scope of the initial search query. Nothing in Copperman would lead one of ordinary skill to receive, from a user, requests for information that include a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept.

In further detail, Copperman describes that documents or other content can be mapped to a knowledge map 200. *See, e.g., Copperman*, para. [0037]. Knowledge map 200 is a collection of taxonomies 210 in which concept nodes are connected by edges. *Id.* The documents in a corpus are assigned to particular concept nodes by a document classifier 445. *See, e.g., id.*, para. [0049].

Rather than matching the terms in a search query directly to these mapped documents, Copperman describes an intermediate step in which certain terms from a user search query are "autocontextualized" by an autocontextualization engine 525. Autocontextualization engine 525 matches text words or phrases from a user search query to concept nodes in a schema. *See, e.g., id.*, para. [0052]. Thus, topics in a search query are "spotted" based on the terms in the search query and assigned to concept nodes. *See, e.g., id.*, para. [0051]. The assigned concepts/topics

are then matched to the mapped content, often through an iterative, guided process. *See, e.g., id.*, para. [0054]. *See also id.*, para. [0051] (describing that terms in a user query that are evidence of concepts are first extracted and then used as the basis for guided search routines).

Copperman's FIGS. 9A-9E illustrate one example of such an iterative, guided search process. *See, e.g., id.*, paras. [0081], [0017]-[0021]. As shown in, e.g., FIG. 9A, a textual search query is first received from a user. *See, e.g., id.*, para. [0081]. As shown in, e.g., FIG. 9C, the response to the search query can include results that match a "primary group feature" that was "spotted" in the terms of the search query, as well as "related features." As best understood, Copperman's features are terms or phrases. *See id.*, para. [0052]. The "related features" can be used to narrow the scope of the search and reduce the number of documents "in play." *See, e.g., id.*, paras. [0054], [0081]. In some instances, the text of a user query can match multiple "primary group features." *See, e.g., id.*, para. [0077].

Thus, like Woods, Copperman's iterative, guided search process also begins with the receipt of a textual search query. As in Woods, such a textual search query is not a request for information from a user that include a definition of a concept list. Although Copperman describes that such textual search query can be autocontextualized, whatever topics or concepts that are spotted during such an autocontextualization are not received from a user. Rather, they are automatically generated by an autocontextualization engine 525.

Also, the "related features" received during Copperman's iterative, guided search process are not requests for information that include a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept,

as recited. To begin with, the terms of the initial search query are already assigned to concept nodes during "autocontextualization." Hence subsequent user selections of related features are not part of a concept list definition that includes those initial search query terms.

Further, there is no reason to believe that the user-selected "related features" are, by themselves, definitions of a concept list. As discussed above, as best understood, Copperman's features are terms or phrases. *See id.*, para. [0052]. There is no reason to believe that the "related features" (e.g., related terms or phrases) are part of a concept list definition as recited in claims 1 and 15. Instead, the user selected "related features" also appear to be stand-alone terms or phrases that simply narrow the result set that is responsive to the search.

Finally, even if one of ordinary skill were to consider Copperman's related features to be part of a single definition that included the terms of the initial search query, this single definition would still not be a concept list definition. In particular, the recited concept list definition includes an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept. However, there is nothing in Copperman that describes that the strength of a relationship between an origin concept and an evaluated concept is somehow received in conjunction with the related features. Rather, the related features are understood to be received as shown, e.g., in Copperman's FIGS. 9A-9E.

The rejection has consistently ignored the actual content of the requests for information in Copperman and consistently ignored the fact that Copperman's knowledge map 200 is not in such requests for information. For example, the Advisory Action mailed March 24, 2008 contends that Cooperman's para. [0132] describes requests for information from a user that include a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the

relationship between the origin concept and the evaluated concept. *See Advisory Action mailed March 24, 2008*, page 2, fourth, fifth, sixth, seventh, and ninth paragraphs. *See also Office action mailed January 10, 2008*, paragraph bridging pages 6-7 (pointing to Copperman's FIG. 2).

Applicant respectfully disagrees. For the sake of convenience Cooperman's para. [0132] and FIG. 9D are now reproduced. FIG. 9D is an example of a portion of a user interface as displayed at a juncture during an illustrative user interaction session. *See id.*, para. [0020]

> "[0132] In a further example, the ranking and/or display of related features 835 for selection by the user is based on the number of times that previous users selected a particular feature choice 835 within the same or similar session context (e.g., with the same or similar confirmed concept nodes deemed relevant to the user query). As an illustrative example, suppose that "TCP-IP" is offered as a related feature 835 in a user session where the Symptom concept node "can't connect" and the Object concept node "network" have already been confirmed as relevant to the user query. In this example, the ranking of 'TCP-IP' with respect to other displayed related features 835 is based on how often previous users selected the various related features when 'can't connect' and 'network' were already confirmed as concept nodes deemed relevant to the user session. In one implementation, each related feature, such as 'TCP-IP', includes a list of confirmed concept nodes with which it has been previously presented. Each such confirmed concept node includes an weight or other indicator including information about how often the particular related feature was selected together with that particular confirmed concept node. For example, the related feature 'TCP-IP' would include a weight for 'can't connect' and 'TCP-IP,' another weight for 'network' and 'TCP-IP', and similar weights for the other confirmed concept nodes with which the 'TCP-IP' related feature 835 has previously been presented. In this example, the ranking and/or display of the 'TCP-IP' related feature 835 is based on such weights. Further description of suitable use-based ranking techniques are described in the above-incorporated Copperman et al. U.S. patent application Ser. No. 09/944,636." *See id.*, para. [0132] (emphasis added).

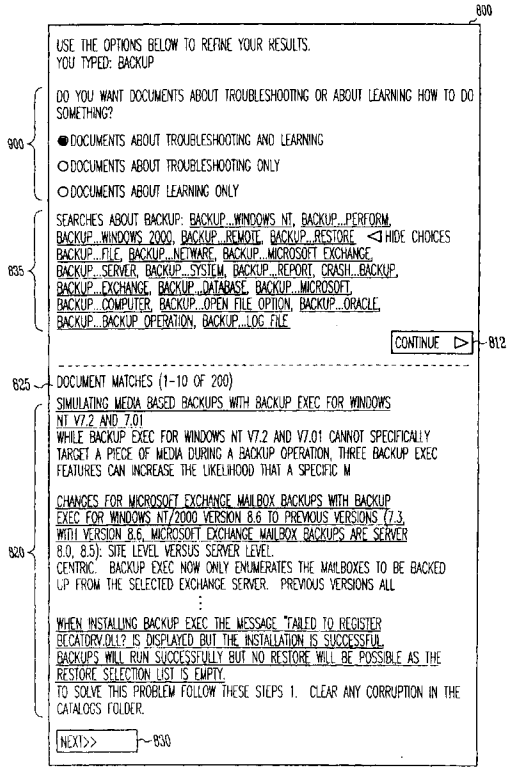

Fig. 9D

As can be seen, the portion of the user interface shown in FIG. 9D does not include any mechanism for receiving requests for information from a user that include a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept. Instead, the user can select a feature choice 835 from a list. It is self-evident that a feature choice is not a definition of a concept list that comprises an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept. Instead, it is the ranking and/or display of the related features 835 that is based on weights that include information about how often a particular related feature was selected together with a particular confirmed concept node. *See id.*, para. [0132].

Thus, Woods and Copperman fail to describe or suggest receiving requests for information, or responding to such requests, as those requests are recited in claims 1 and 15. Even if Woods and Copperman were combined, one of ordinary skill would not arrive at the recited subject matter. Accordingly, claims 1 and 15 are not obvious over Woods and Copperman on this basis.

Further, the rejections of claims 1 and 15 are based on the contentions that one of ordinary skill would find it obvious to combine Woods and Copperman to arrive at the recited subject matter "in order to classify documents according to the most pertinent concept or concepts." *See Office action mailed January 10, 2008*, page 4, line 1-4; Advisory Action mailed March 24, 2008, page 2, para. 8. It appears that this contention is based on Woods' and Copperman's classification of documents based on semantic networks. As discussed above, Woods describes that a corpus of documents is to be analyzed to construct term/concept relationship network 110. Copperman describes that documents in a corpus are to be assigned to particular concept nodes in a knowledge map 200. Thus, in both cases, Woods and Copperman are concerned with classifying documents.

However, claims 1 and 15 recite that a request for information includes a definition of a concept list that comprises an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept. Claims 1 and 15 also recite that a target score that represents how well the document regions relate to the concept list is determined and that an extract is output in response to the request for information based on the target score.

These features are not part of the classification of documents, nor is there any reason to believe that those of ordinary skill would have found these features obvious based on the classification of documents in Woods and Copperman. Indeed, as discussed above, given that the classification of documents in Woods and Copperman is tangential to the receipt and response to requests for information recited in claims 1 and 15, even if Woods and Copperman were combined "in order to classify documents," there is no reason to believe that those of ordinary skill would arrive at the recited subject matter.

For these and other reasons, claims 1 and 15 are not obvious over Woods and Copperman. Applicant respectfully requests that the rejections of claims 1, 15, and the claims dependent therefrom be withdrawn.

Ground 2: Rejections under 35 U.S.C. § 103(a) over Woods, Copperman, and Sacco

Claim 47 was rejected under 35 U.S.C. § 103(a) as obvious over Woods, Copperman, and Sacco.

As discussed above, claim 47 relates to a computer-implemented method for extracting a subset of a document. The method includes receiving, from a user, a request for information that describes a combination of two or more concept lists, receiving a description of a document region targeted for extraction, accessing a document, based on the target definition and the document regions targeted for extraction, extracting one or more regions of the accessed document, and making the extracted regions available for output in response to the request for information. Each concept list is defined by an origin concept, a relationship between the origin

concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept. The two or more concept lists are combined using an operation to define a target definition that is to be detected.

As discussed above, Woods and Copperman neither describe nor suggest receiving a request for information from a user that includes a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept. Further, nothing in Woods and Copperman would make it obvious for one of ordinary skill to receive such a definition of a concept list, especially when the nature of Woods' single search query phrase and Copperman's initial search query and any subsequently selected "related features" are considered.

Applicant thus submits that Woods and Copperman also fail to describe or suggest receiving, from a user, a request for information that describes a combination of two or more concept lists, where each concept list is defined by an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept.

Sacco does nothing to remedy these deficiencies in Woods and Copperman. Instead, Sacco is understood to describe an iterative, taxonomy-driven process for browsing and retrieving information in large heterogeneous databases. *See, e.g., Sacco*, col. 1, line 7-10. In Sacco, a set of documents is classified under a taxonomy. *See, e.g., id.*, col. 2, line 50-53. Sacco's taxonomy "is usually a tree, but lattices (deriving from a concept having more than one father) are allowed." *See, e.g., id.*, col. 2, line 60-61.

The taxonomy can be displayed for a user with father-to-son relations and son-to-father relations (i.e., lists of sons and fathers) of each concept. *See, e.g., id.*, col. 3, line 21. A user can select concepts in the taxonomy, as well as Boolean operations for combining them. *See, e.g., id.*, col. 3, line 22-25. This process can be iterative (i.e., these steps can be repeated). *See, e.g., id.*, col. 3, line 27-28. A user "will then be presented with a reduced taxonomy for the selected set of documents, which can be iteratively further refined." *See, e.g., id.*, col. 2, line 14-16.

While Sacco does indeed receive requests for information from a user that are not limited tp the terms in Wood's search query phrases and Copperman's initial search query and any subsequently selected "related features," Sacco does not render requests for information that describes a combination of two or more concept lists obvious to those of ordinary skill. To begin with, users in Sacco select concepts in a taxonomy that has been used to classify documents. Such a selection of concepts neither describes nor suggests that concept lists—defined by an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept—are received in a request for information. Instead, a father or son concepts (and Boolean operation) are selected by Sacco's users.

Thus, even if Woods, Copperman, and Sacco were combined, one of ordinary skill would not arrive at the recited subject matter. Accordingly, claims 1 and 15 are not obvious over Woods, Copperman, and Sacco. Applicant respectfully requests that the rejections of claim 47, and the claims dependent therefrom, be withdrawn.

Please apply the $510 brief fee and one-month extension of time and any other charges or credits to Deposit Account No. 06-1050.

Respectfully submitted,

Date: July 7, 2008

John F. Conroy
Reg. No. 45,485

Fish & Richardson P.C.
12390 El Camino Real
San Diego, California 92130
Telephone: (858) 678-5070
Facsimile: (877) 769-7945

JFC/jhg
10839624.doc

Appendix of Claims

1. A computer-implemented method, comprising:

receiving, from a user, a request for information that includes a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept, and a target scope that characterizes a document region to which the concept list is to be applied;

receiving a definition of an extraction rule, wherein the extraction rule definition comprises an extraction scope that characterizes a document region to be extracted;

determining a target score for the document regions of the article, wherein the score represents how well the document regions relate to the concept list;

applying the extraction rule to the article to determine an extract from the article, wherein the application of the extraction rule is based on the determined target score; and

outputting the extract in response to the request for information.

2. The method of claim 1, wherein applying the extraction rule comprises extracting a plurality of extracts.

3. The method of claim 2, further comprising sorting the extracts based on the extraction rule.

4. The method of claim 1, further comprising selecting a first extract from the article for output based on the target score.

5. The method of claim 1, further comprising receiving a target score formula for determining the target score.

6. The method of claim 5, wherein determining the target score comprises using the target score formula.

7. The method of claim 1, wherein the request for information further comprises a gist defined as a vector of weighted concepts.

8. The method of claim 1, wherein the request for information further comprises a concept set that comprises the concept list and a second concept.

9. The method of claim 8, wherein the second concept comprises a product of set operations on two or more other concepts.

10. The method of claim 7, wherein the gist comprises a user defined gist.

11. The method of claim 7, wherein the gist comprises a calculated gist of a document region.

12. The method of claim 1, wherein the document region characterized by the target scope comprises an article, a sentence, or a term.

13. The method of claim 1, wherein the document region characterized by the extraction scope comprises an article, a sentence, or a term.

14. The method of claim 1, further comprising preprocessing the article, wherein preprocessing comprises:

determining at least one concept associated with the article; and

determining a gist of the article.

15. An article comprising one or more computer-readable data storage media containing program code operable to cause one or more machines to perform operations, the operations comprising:

receiving, from a user, a request for information that includes a definition of a concept list comprising an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept, and a target scope that characterizes a document region to which the target rule is to be applied;

receiving a definition of an extraction rule, wherein the extraction rule definition comprises an extraction scope that characterizes a document region to be extracted;

determining a target score for the document regions of the article, wherein the score represents how well the document regions relate to the concept list;

applying the extraction rule to the article to determine an extract from the article, wherein the application of the extraction rule is based on the determined target score; and

outputting the extract in response to the request for information.

16. The article of claim 15, wherein applying the extraction rule comprises extracting a plurality of extracts..

17. The article of claim 16, further comprising sorting the extracts based on the extraction rule.

18. The article of claim 15, further comprising selecting a first extract from the article for output based on the target score.

19. The article of claim 15, further comprising receiving a target score formula for determining the target score.

20. The article of claim 19, wherein determining the target score comprises using the target score formula.

21. The article of claim 15, wherein the request for information further comprises a gist defined as a vector of weighted concepts.

22. The article of claim 15, wherein the request for information further comprises a concept set that comprises the concept list and a second concept.

23. The article of claim 22, wherein the second concept comprises a product of set operations on two or more other concepts.

24. The article of claim 21, wherein the gist comprises a user defined gist.

25. The article of claim 21 wherein the gist comprises a calculated gist of a document region.

26. The article of claim 15, wherein the document region characterized by the target scope comprises an article, a sentence, or a term.

27. The article of claim 15, wherein the document region characterized by the extraction scope comprises an article, a sentence, or a term.

28. The article of claim 15, further comprising preprocessing the article, wherein preprocessing comprises:

determining at least one concept associated with the article; and

determining a gist of the article.

29. The method of claim 1, wherein:

the origin concept comprises a group of related words, relationships with other concepts, the strengths of the relationships, and statistics regarding the usage of the origin concept in a language.

Claims 30.-34. (Canceled)

35. The method of claim 1, wherein receiving the definition of the extraction rule further comprises receiving a definition of a sort order in which extracts are to be sorted for output.

36. The article of claim 15, wherein receiving the definition of the extraction rule further comprises receiving a definition of a sort order in which extracts are to be sorted for output.

37. The method of claim 1, wherein the distance comprises a numeric representation of the strength of the relationship between the origin concept and the evaluated concept.

38. The method of claim 1, wherein the relationship comprises one of "is a product of", "is a part of", "is a kind of", "has kind", or "has part".

39. The method of claim 1, wherein the origin concept comprises at least one search term.

40. The method of claim 8, wherein the concept set further comprises at least one set operation.

41. The method of claim 40, wherein the set operation comprises one of "AND", "OR", and "AND NOT".

42. The article of claim 15, wherein the distance comprises a numeric representation of the strength of the relationship between the origin concept and the evaluated concept.

43. The article of claim 15, wherein the relationship comprises one of "is a product of", "is a part of", "is a kind of", "has kind", or "has part".

44. The article of claim 15, wherein the origin concept comprises at least one search term.

45. The article of claim 22, wherein the concept set further comprises at least one set operation.

46. The article of claim 45, wherein the set operation comprises one of "AND", "OR", and "AND NOT".

47. A computer-implemented method for extracting a subset of a document, comprising:

receiving, from a user, a request for information that describes a combination of two or more concept lists, wherein each concept list is defined by an origin concept, a relationship between the origin concept and an evaluated concept, and a distance representing a strength of the relationship between the origin concept and the evaluated concept , wherein the two or more concept lists are combined using an operation to define a target definition that is to be detected;

receiving a description of a document region targeted for extraction;

accessing a document;

based on the target definition and the document regions targeted for extraction, extracting one or more regions of the accessed document; and

making the extracted regions available for output in response to the request for information.

48. The method of claim 47, wherein the origin concepts each comprise a lexical concept defined by a group of related words and relationships with related concepts.

Evidence Appendix

None.

Related Proceedings Appendix

None.